illumina®

ForenSeq™ Universal Analysis Software

Interrogate the broadest range of forensically relevant loci with a simple user interface and powerful analysis algorithms.

Highlights

- **Complete, Sample-to-Answer Workflow**
  Manage the sequencing workflow from run set-up to data analysis and report generation
- **Perform Automated Data Visualization and Reporting with Simple Graphical User Interface**
  Compare samples automatically, generate population statistics, and view data through an intuitive user interface
- **Generate Investigative Leads Through Analysis Software Advances**
  Estimate visible traits such as hair color and eye color as well as biogeographical ancestry

Introduction

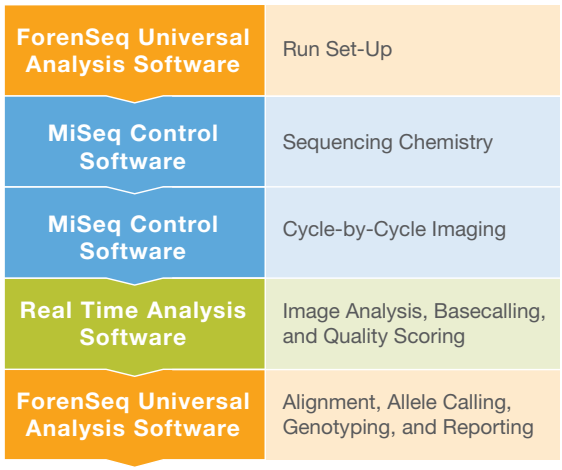
Illumina ForenSeq Universal Analysis Software is part of the MiSeq FGx™ Forensic Genomics System—a fully validated[1] next-generation sequencing (NGS) system specifically designed for use in forensic genomics applications. The MiSeq FGx System and ForenSeq Universal Analysis Software enable the simultaneous analysis of approximately 200 forensically relevant single nucleotide polymorphisms (SNPs) and short tandem repeats (STRs)—including marker sets not routinely available with traditional methods (Table 1). From sequencing library* and file management to complete data analysis and reporting, the analysis software provides the full functionality needed to analyze and interpret the widest range of casework and database samples.

To maximize flexibility and scalability of the MiSeq FGx System for each laboratory, the ForenSeq Universal Analysis Software is sold separately on a standalone, dedicated server.

Complete, Sample-to-Answer Workflow

The MiSeq FGx Forensic Genomics System uses a pipeline of software applications to perform sequencing runs and complete data analysis (Figure 1). ForenSeq Universal Analysis Software guides the sequencing workflow from run-setup and sample data entry to the final stages of data analysis and report generation. MiSeq FGx Control Software (MCS), installed on the MiSeq FGx instrument control computer, captures flow cell images, operates the flow cell stage, and controls reagent delivery and temperature. During the run, Real-Time Analysis (RTA) software, performs image analysis, base calling, and assigns base-by-base quality scores. ForenSeq Universal Analysis Software then initiates the final stages of analysis, including demultiplexing, sequence alignment, allele calling, genotyping, and reporting.

Figure 1: ForenSeq Universal Analysis Software Workflow



The MiSeq FGx Forensic Genomics System uses the MiSeq FGx Control Software, Real-Time Analysis, and ForenSeq Universal Analysis Software to operate the sequencing instrument and perform data analysis.

Table 1: Simultaneous Analysis of Forensically Relevant Loci with ForenSeq Universal Analysis Software

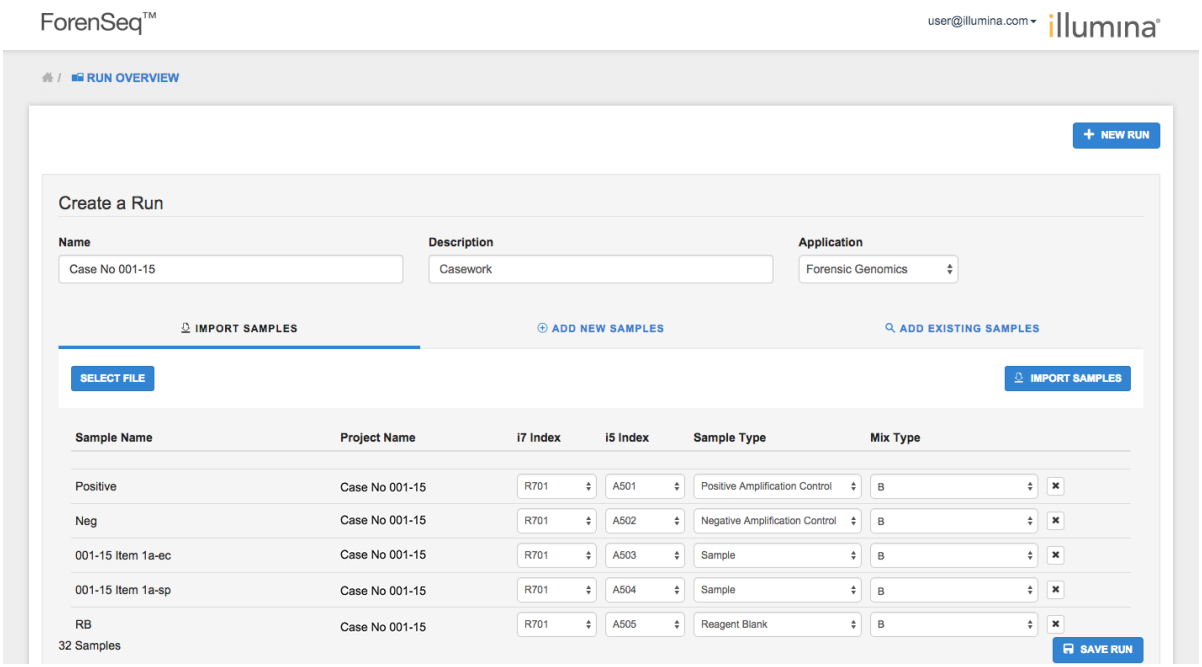
Feature	Number of Markers[a]
Global Autosomal STRs	27
Y-STRs	24
X-STRs	7
Identity SNPs	94
Phenotypic SNPs	22
Biogeographical Ancestry SNPs	56
Total Number of Loci [b]	> 200

a. SNP and STR chromosome locations can be found in the ForenSeq DNA Signature Prep Kit User Guide (support.illumina.com/downloads/forenseq-dna-signature-prep-guide-15049528.html).

b. Over 200 markers analyzed when running primer set B.

* A sequencing "library" is a collection of amplified DNA fragments from a single DNA sample.

Figure 2: ForenSeq Universal Analysis Run Overview Screen



The Run Overview screen features easy data entry fields for Run Name and Run Description. Sample information, index combinations, and application options can also be imported with a tab delimited .txt file. Sample types include Known Sample, Forensic Sample, Reagent Blank, Negative Amplification Control, Positive Amplification Control, and Human Sequencing Control.

Perform Automated Data Visualization and Reporting with Simple Graphical User Interface

Illumina ForenSeq Universal Analysis Software provides an easy data analysis workflow and user-friendly viewing and reporting features. The graphical user interface has a simple, intuitive design that can be viewed from any laboratory computer through the Google Chrome web browser. The Run Overview screen features easy run-setup, sample information, and index tracking options (Figure 2). Within the analysis software, sequencing run data can be associated with specific, user-defined projects and results can be viewed by run or by project in the Projects screen.

At run completion, ForenSeq Universal Analysis Software offers data visualization through several types of display screens. The Sample Details screen provides a summarized view of SNP and STR locus intensity scatter plots. Below the intensity charts, individual allele calls are displayed, along with convenient icons displayed when specific quality flags are triggered (Figure 3). To view more indepth reports on specific loci, drill down using the Locus Detail pop-up screens. The Locus Detail screens display intensity charts, the base-by-base DNA target sequence, stutter, and more (Figure 4). The analysis software also offers a full suite of run metrics and sample quality controls that can be evaluated after the run. Quality metrics and sample quality controls are displayed with color indicators allowing forensic analysts to quickly scan the results and determine whether controls fall within the recommended ranges.

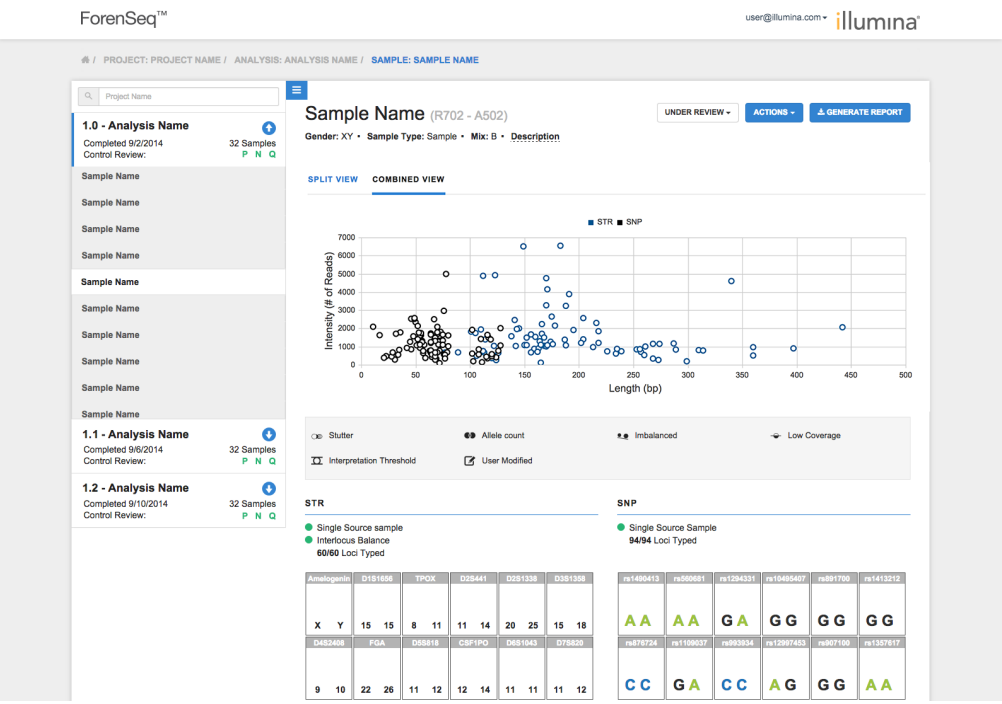
Project Detail reports and Sample Detail reports can be generated automatically within the software. Each report can be easily printed or exported in .xlsx format. Password protected user accounts ensure secure access to the MiSeq FGx sequencing instrument and to the analysis server.

Generate Investigative Leads through Analysis Software Advances

Compared to current capillary electrophoresis-based methods, Illumina ForenSeq Universal Analysis Software provides a number of data analysis advantages. These include, the capacity to interrogate a greater number of markers (Table 1) and the ability to recover the maximum amount of useful genomic information from degraded DNA, low quality DNA, or complex mixtures[2]. Furthermore, all markers, including SNPs, autosomal, X-, and Y-STRs, can be analyzed simultaneously using a single, streamlined workflow.

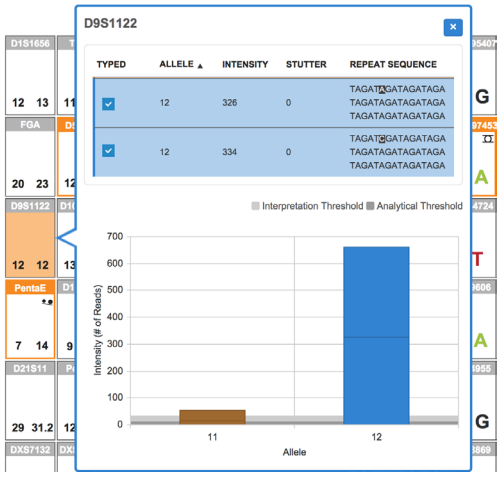
Beyond analysis of autosomal, X-, and Y-STRs, the software enables the analysis of marker sets not routinely available with traditional capillary electrophoresis methods. These include a dense set of identity-informative SNPs (iiSNPs)[3,4], which are informative for source attribution, phenotypic-informative SNPs (piSNPs)[5], which provide estimates of eye color (blue, intermediate, brown) and hair color (brown, red, black, blond), and biogeographical ancestry-informative SNPs (aiSNPs)[6]. Biogeographical ancestry estimation is presented as a principal component analysis (PCA) plot relative to major population

Figure 3: ForenSeq Universal Analysis Software Sample Details Screen



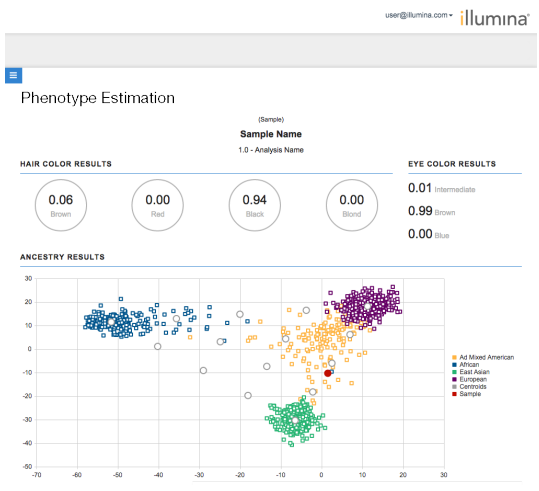
The Sample Details screen displays overall sample results as an intensity plot, along with detailed summary tables for all loci included in the sequencing library.

Figure 4: ForenSeq Universal Analysis Software Locus Detail Screen



The pop-up Locus Detail screen shows the allele call, intensity bar graph, and stutter. View the Repeat Sequence column to see the full, base-by-base target sequence, and any intra-STR variation present in the sample.

Figure 5: ForenSeq Universal Analysis Software Estimation of Visible Traits



The ForenSeq Universal Analysis Software enables estimation of visible traits including eye and hair color. Markers for biogeographical ancestry can also be tested and reviewed.

AAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAA
AATCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACC
AACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAA
TTAACGTACCATTAAGAGCTACCGTGCAACGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATA
AAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAA
AAGATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGA
AACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAG

Table 2: ForenSeq Server Components and Specifications

4 × Seagate 1 TB Sata hard drives
RAID controller allowing for data redundancy and speed
Intel 2 GHz × 64 processor with 6 cores/12 threads
Intel socket R server board
32 GB DDR3 RAM
550 W power supply
ULL, FCC, CE certified
Windows Server 2012 R2 Standard with 5 CALs

groups[7] (Figure 5). Centroids within the plots then provide logical groupings that give meaningful context to the unknown sample.

Additionally, ForenSeq Universal Analysis Software can detect intra-STR sequence variation that discloses additional identity informative data beyond length variation (Figure 4). These data can be important in kinship analyses and mixture analysis, where DNA contributors share equi-length alleles that can only be distinguished by comparing their base-by-base sequences.

Summary

ForenSeq Universal Analysis Software contains comprehensive sample management and analytical capabilities, including sample and index management, application-specific workflows, data visualization at sample and locus levels, quality flags to simplify data analysis, and easily exportable reports.

Learn More

To learn more about ForenSeq Universal Analysis Software, visit: www.illumina.com/informatics/sequencing-microarray-data-analysis/forenseq-universal-analysis-software.ilmn

Ordering Information

Product	Catalog No.
ForenSeq Universal Analysis Software and Server	SE–550–1001

To learn about the MiSeq FGx instrument, visit: www.illumina.com/systems/miseq-fgx.ilmn

To learn about the ForenSeq DNA Signature Prep Kit, visit: www.illumina.com/products/forenseq-dna-signature-kit.ilmn

To learn about mitochondrial DNA analysis with the MiSeq FGx System, visit: www.illumina.com/applications/forensics/missing_persons_and_unidentified_human_remains.ilmn

References

1. The full MiSeq FGx System workflow is validated per the Scientific Working Group on DNA Analysis Methods (SWGDAM) guidelines (www.swgdam.org).
2. Illumina (2014) ForenSeq DNA Signature Prep Kit Data Sheet (www.illumina.com/products/forenseq-dna-signature-kit.ilmn).
3. Kidd KK, Pakstis AJ, Speed WC, Grigorenko EL, Kajuna SL, et al. (2006) Developing a SNP panel for forensic identification of individuals. Forensic Sci Int 164(1): 20–32.
4. Sanchez JJ, Phillips C, Børsting C, Balogh K, Bogus M, et al. (2006) A multiplex assay with 52 single nucleotide polymorphisms for human identification. Electrophoresis 27(9): 1713–1724.
5. Walsh S, Liu F, Wollstein A, Kovatsi L, Ralf A, et al. (2013) The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci Int Genet 7(1): 98–115.
6. Kidd KK, Speed WC, Pakstis AJ, Furtado MR, Fang R, et al. (2013) Progress toward an efficient panel of SNPs for ancestry inference. Forensic Sci Int Genet 10: 23–32.
7. Phillips C, Prieto L, Fondevila M, Salas A, Gómez-Tato A, et al. (2009) Ancestry analysis in the 11-M Madrid bomb attack investigation. PLoS One 4(8): e6583.

Illumina • 1.800.809.4566 toll-free (U.S.) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

FOR RESEARCH, FORENSIC, OR PATERNITY USE ONLY – NOT FOR ANY CLINICAL OR THERAPEUTIC USE IN HUMANS OR ANIMALS.

© 2015 Illumina, Inc. All rights reserved.
Illumina, ForenSeq, MiSeq FGx, the pumpkin orange color, and the streaming bases design are trademarks of Illumina, Inc. and/or its affiliates in the U.S. and/or other countries. All other names, logos, and other trademarks are the property of their respective owners.
Pub. No. 1470-2014-002 Current as of 17 June 2015

illumina®

AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
TCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTT
CGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACC
AACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAA
AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
GATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGAGC
CGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAGTA